Some Considerations on Six Degrees of Separation from A Theoretical Point of View

Norihito Toyota

Hokkaido Information University, Ebetsu, Nisinopporo 59-2, Japan

email :toyota@do-johodai.ac.jp

Abstract

In this article we discuss six degrees of separation, which has been proposed by Milgram, from a theoretical point of view. Simply if one has k friends, the number N of indirect friends goes up to $\sim k^d$ in d degrees of separation. So it would easily come up to population of whole world. That, however, is unacceptable. Mainly because of nonzero clustering coefficient C, N does not become $\sim k^d$. In this article, we first discuss relations between six degrees of separation and the clustering coefficient in the small world network proposed by Watt and Strogatz[2],[3]. Especially, conditions that $(N) >$ (population of U.S.A or of the whole world) arises in the WS model is explored by theoretical and numerical points of view. Secondly we introduce an index that represents velocity of propagation to the number of friends and obtain an analytical formula for it as a function of C, K, which is an average degree over all nodes, and some parameter P concerned with network topology. Finally the index is calculated numerically to study the relation between C, K and P and N.

keywords: Six Degrees of separation, Small world Network, Propagation Coefficient, Watt-Strogatz Model, Clustering Coefficient, Average Path Length

1   Introduction

In 1967, Milgram made a great impact on the world by advocating the concept "six degrees of separation" by an social experiment in a celebrated paper [1]. "Six degrees of separation" shows that people have a narrow circle of acquaintances. A series of social experiments made by him suggest that all people in USA are connected through about 6 intermediate acquaintances. In this paper we inspect from a rather theoretical point of view that this phenomenon, so called "six degrees of separation" is not so surprising and if anything natural one.

This article is first motivated by a following simple consideration; If every person has K acquaintances, so that after L steps of intermediate acquaintances the person would be able to convey a mail, in general information, to about $S = K^L$ persons. With the proviso, however, that the network of acquaintances has a tree structure without any loop (no clustering coefficient), evaluating it in more detail, it is

$$S = \sum_{i=0}^{L-1} K(K-1)^i = \frac{K(K^{L-1}-1)}{K-1}, \tag{1}$$

where it is assumed that the relation of acquaintances is symmetric. Thus information will spread out among exponentially many persons from one person with L steps. Though a person that received information, of course, may convey the information to only a part of his/her acquaintances, when a network has a tree structure, six degrees of separation for any two persons would not be so mysterious.

Real networks, however, naturally have structures with loops. If there is some loops or an effective clustering coefficient in the network of acquaintances, this discussion will greatly altered. One of aims of this article is to evaluate the effect of clustering coefficient on propagation of information. How much does the clustering coefficient reduce the population that information is provided ? We consider it from three points of view. They consist of the following three;

1. To inspect six degrees of separation for Watts-Strogatz type model where analytic expressions for the clustering coefficient and the average path length are found.

2. To inspect six degrees of separation for more general small world networks with uniform clustering coefficient.

3. To consider empirical networks such as a network in Mixi and a network of actor/actress besed on data of Bacon game.

The plan of this article is as follows. In the next section we argue on Watts-Strogatz type small world networks[2, 3] where theoretical evaluation of the average length and the clustering coefficient has been made. We study the propagation of information on the networks by making numerical analyses based on these formulae, To study more general small world networks, we present a propagation coefficient model in the section 3. We can not analyze, however, large class of small world networks including scale free networks due to technical reasons at present. We adopt a homogeneous hypothesis which would be explained in detail in the section 3. Various types of homogeneity are postulated in the propagation coefficient model. In the section 4, we analyze experimental data, such as a network in Mixi and a network based on data of Bacon game, as to estimate clustering coefficient. Concluding remarks are given in the last section.

2 Watts-Strogatz Model

In this section we investigate the effect of the clustering coefficient on diffusion of information in Watts-Strogatz type small world networks. The analytic expressions for the clustering coefficient and the average path length in the networks have been found. For the original Watts-Strogatz version of the model, the clustering coefficient $C(p)$ has been given [18] by

$$C(p) = \frac{3K - 6}{4K - 2}(1 - p)^3, \tag{2}$$

where n is the network size, that is, the number of nodes on the network, K is the average degree and p is the rewiring probability. Moreover the average node-node separation $L(p)$ in a modified version of the original Watts-Strogatz model has found in the limit of low density of shortcuts[19];

$$L(p) = \frac{2n}{K}F(\frac{nKp}{2}), \quad F(x) = \frac{1}{2\sqrt{x^2 + 2x}}\tanh^{-1}\sqrt{\frac{x}{x + 2}}, \quad \text{for small } p. \tag{3}$$

In the modified version of the original small world model in which shortcuts edges are added between randomly chosen node pairs, no bonds are removed. Here edges are not rewired, rather adding edges thus ensuring that the modified network stays connected. We attempt analyses of six degrees of separation based on these two formulae.
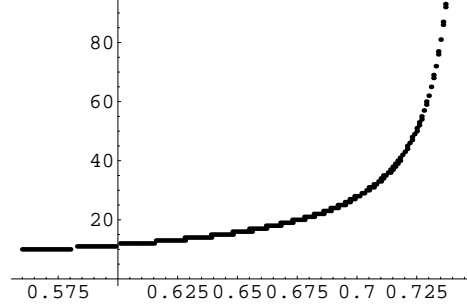
2.1 Parameter Regions for Six Degrees of Separation

In this subsection, we explore possible regions of parameters introduced in the previous section for six degrees of separation. First we explore parameters' regions in order that information can spread among on average 10^9 persons by 6 ± 1 steps. Taking $L = 6 \pm 1$ and $n = 10^9$ in Eq. (3), we explore the possible regions in a $p - K$ space. We can numerically find them within meaningful values of the parameters. They are shown in Fig.1. As far as p is not extremely small that is the small world region, that K, the average number of contacts of a person, is several tens is sufficient for information to spresd from one person to 10^9 persons.

The solution for $C - K$ space where p is eliminated from Eq.(1) and Eq.(2) is described in Fig. 2. Though Eq. (3) only holds only at small p, then at least we can work out two equations simultaneously. Since as p becomes large, C becomes small, the validity of the analysis would be lost for small C. Fig. 2 asserts that it is sufficient that K is several tens for relatively large values of C.
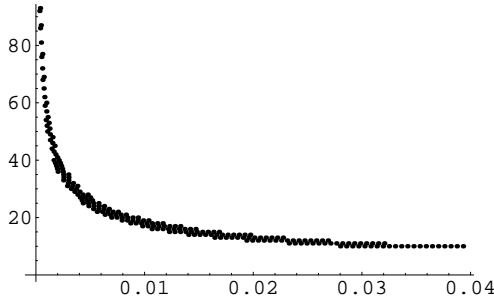
FigD1. p-K plots in SW model.   FigD2. C-K plots in SW model.

@

Through these analyses, we conclude that following regions of parameters are roughly needed;

$$C = 0.55 \sim 0.75,$$
$$p = 0.01 \sim 0.04,$$
$$K = 15 \sim 25,$$

in order that information can spread from one person to 10^9 persons at about 6 steps on Watts-Strogatz type small world networks.

2.2 Population propagated at $L = 6$

In this subsection we consider that to how many persons information can propagate at exact six steps from a person. We find adequate regions in a parameters' space $K - C$ when n falls between 10^7 and 10^9 at $L = 6$. For each K, typical values of C are listed in Table 1.

Table 1. C and K for $n \sim 10^7 \sim 10^9$

K	9	10	12	15	20	25	49	100	147	194
C	0.54	0.57	0.61	0.64	0.68	0.69	0.724	0.738	0.74	0.744

As K becomes large, n obviously becomes large. From Table 1, K is not so large even for rather large C. Thus information can readily spread to about a billion persons at 6 steps, if C took fairly small value.

Table 2. Empirical Networks with large $C > 0.5$

Network	Number of Vertices	$< K >$	C	$< L >$	Index
Company directors	7673	14.4	0.59		
Coauthorships in the SPIREES e-archive	56627		0.726	4	1.2
Collaboration net collected from math.journals	70975		0.59	2.1294	9.5
Collaboration net collected from neurosci.journals	209293		0.76	2.4	6
metabolic network	315		0.59		
World Web	470000		0.69/0.44	1.5/2.7	2.65

So far we have seen that it has been possible to realize six degrees of separation even rather large C. There, however, are not so many empirical networks with large C. Networks with rather large C that have discovered so far are listed in Table 2 [19],[14],[15],[16],[14],[10]. Here blanks represent that they are unknown and the index in the rightest column is that for scale free. These networks mostly have scale free nature, so that they are not Watts-Strogatz type small world networks.

3 Propagation Coefficient Model

We study Milgram-like-propagation of information in wider class of networks, including those other than small world networks. First we focus our attention on any one node, which is the node in

0-generation, in general networks. Next we explore all nodes connected with the first node, which are nodes in the 1-generation. Next we explore all nodes connected with the nodes in the 1-generation apart from the nodes of 0-generation, which are nodes in the 2-generation. We continue these procedures until all nodes on a network are covered by these procedures. These procedures are effective in any complex networks. There are only two generations in complete graphs. The maximal generation number N_G of a network is larger than the diameter of the network by 1 and so the maximal generation number G is just equal to the diameter. This picture of networks make the analyses of the propagation of information on the network manageable.

Now we introduce some geometrical quantities that used in this article together with their notation.

G_i means the i-th generation and n_i is the numbers of generation G_i.

N is the total number of nodes of a network or the size of the network.

$k_{i,i+1}^{(j)}$ is the number of edges from node j in G_i to nodes in G_{i+1}.

$k_{i,i}$ is the number of edges connected between the same generation G_i.

C_i is a contribution to the clustering coefficient produced by edges in G_i.

$K^{(j)}$ is the degree of the node j.

We define the propagation coefficient from G_i to G_{i+1} by $\overline{k}_{i,i+1}$ as an average of $k_{i,i+1}$.

We make the assumptions for simplicity of analyses.

1. The size of the network is infinite.
2. A parameter q_j is the possibility that a node j has two parents.
3. There is no backflow in the propagation of information.
4. The homogeneous hypothesis ;

$$C_i = \overline{C} = const. \tag{4}$$

$$q_j = q = const. \quad but \quad q = 0 \quad at \quad G_0 \quad generation. \tag{5}$$

$$K^{(j)} = K = const \quad (nearly \quad equal \quad degree). \tag{6}$$

Under these assumptions, the following relations hold;

$$\overline{k}_{i,i+1} = \frac{\sum_{j \in G_i}^{n_i} k_{i,i+1}^{(j)}}{n_i}, \tag{7}$$

$$K = 1 + \frac{2k_{i,i}}{n_i} + \overline{k}_{i,i+1}, \tag{8}$$

$$n_{i+1} = \overline{k}_{i,i+1} n_i (1 - q), \tag{9}$$

$$N = \sum_{i}^{d} n_i. \tag{10}$$

In order to investigate the effect of the clustering coefficient upon the propagation of information on a network, we consider when does the clustering coefficient increase in this picture of networks. Notice that there are two cases that make a contribution to the clustering coefficient in this picture of networks. One is the case that ages are linked together with nodes in the same generation. The other is the case that one node has two parent nodes that are linked each other. They are shown in the Fig.3.

G₀

G₁

G₂
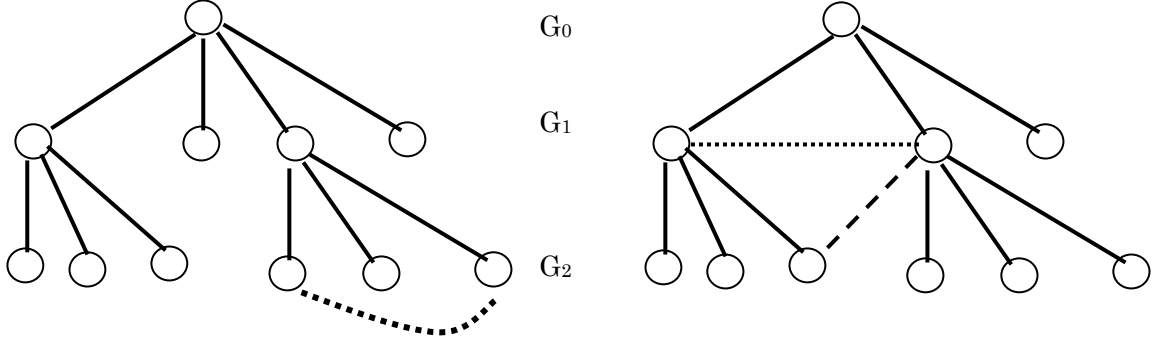
Links between same generations.   The cases that a node has two parents.

FigD3. Two patterns that contribute to a clustering coefficient in G_i generation. The possibility that the right hand pattern occurs is q. @@

@

Since each node has K edges from the assumption 4., C_i is generally obtained by

$$C_i = \frac{t_i(q)}{{}_K C_2}. \tag{11}$$

We investigate two types of contributions to $t_i(q)$ in order.

First we consider the case in the left hand side of Fig.3, that is to say $t_i(0)$. When one edge between the same generation G_i is added, the average probability that the edge is just one between children nodes with a common parent node in G_{i-1} is $\overline{k}_{i-1,i}C_2/n_iC_2$. So as $k_{i,i}$ edges in G_i are totally added, the number of triangles in G_i become

$$t_i(0) = \frac{\overline{k}_{i-1,i}C_2 k_{i,i} n_{i-1}}{{}_{n_i}C_2} = \frac{k_{i,i}(\overline{k}_{i-1,i}-1)}{\overline{k}_{i-1,i}n_{i-1}-1} = \frac{k_{i,i}(\overline{k}_{i-1,i}-1)}{\prod_{l-1}^{i}\overline{k}_{l-1,l}-1}, \tag{12}$$

since there are n_{i-1} families in G_i and $n_i = n_{i-1}\overline{k}_{i-1,i}$ at $q = 0$ in Eq. (9) is used in the last equality. We can obtain explicit expressions for first a few $t_i(0)$;

$$\begin{aligned}
t_1(0) &= k_{1,1}, \\
t_2(0) &= k_{2,2}\frac{\overline{k}_{1,2}-1}{K\overline{k}_{1,2}-1}, \\
t_3(0) &= k_{3,3}\frac{\overline{k}_{2,3}-1}{K\overline{k}_{1,2}\overline{k}_{2,3}-1}, \\
\cdots \quad \cdots & \qquad \cdots.
\end{aligned} \tag{13}$$

In the other case, the possibility that two parents are linked in their generation G_{i-1} is $k_{i-1,i-1}/{}_{n_{i-1}}C_2$ and the number that nodes in G_i have two parent nodes is $qn_{i-1}\overline{k}_{i-1,i}$. Combining these two facts, the contribution $t'_i(q)$ to $t_i(q)$ of this case is given by

$$t'(q) = (qn_{i-1}\overline{k}_{i-1,i})\frac{k_{i-1,i-1}}{{}_{n_{i-1}}C_2} = \frac{2q\overline{k}_{i-1,i}k_{i-1,i-1}}{n_{i-1}-1} = \frac{2q\overline{k}_{i-1,i}k_{i-1,i-1}}{\prod_{l-1}^{i}\overline{k}_{l-1,l}(i-q)^{1-q}-1}, \tag{14}$$

where Eq.(9) is used in the last equation.

Thus $t_i(q)$ is finally obtained by

$$t_i(q) = \frac{1}{{}_{n_i}C_2}\Big(k_{i,i}n_i \times_{\overline{k}_{i-1,i}}C_2\Big) + \frac{1}{{}_{n_{i-1}}C_2}\Big(qk_{i-1,i-1}n_{i-1}\overline{k}_{i-1,i}\Big), \tag{15}$$

where each term in parentheses shows the contribution from the left hand figure and the right one in the Fig.3, respectively.

5

Thus the clustering coefficients C_i in the generation G_i are obtained by

$$C_i = \frac{\overline{k}_{i-i,i}n_{i-1}}{K(K-1)}\left(\frac{(\overline{k}_{i-1,i}-1)(K-1-\overline{k}_{i,i+1})}{\overline{k}_{i-1,i}n_{i-1}-1} + \frac{2q(K-1-\overline{k}_{i-1,i})}{n_{i-1}-1}\right), \tag{16}$$

where the equation derived from Eq.(4)

$$k_{i,i} = \frac{n_i}{2}(K-1-\overline{k}_{i,i+1}) = \frac{n_{i-1}\overline{k}_{i-i,i}}{2}(K-1-\overline{k}_{i,i+1}) \tag{17}$$

is used.

By using $n_0 = 1$ and $\overline{k}_{0,1} = n_1 = K$, we can give explicit expressions for first a few C_i;

$$\begin{aligned}
C_1 &= 1 - \frac{\overline{k}_{1,2}}{K-1}, \\
C_2 &= \frac{\overline{k}_{1,2}}{K-1}\left(\frac{(\overline{k}_{1,2}-1)(K-1-\overline{k}_{2,3})}{K\overline{k}_{1,2}-1} + \frac{2q(K-1-\overline{k}_{2,3})}{K-1}\right), \\
&\cdots \quad \cdots \qquad \cdots.
\end{aligned} \tag{18}$$

Using the homogeneous hypothesis $C_i = \overline{C} = const.$, we can express every $\overline{k}_{i-1,i}$ in terms of the $\overline{C}$. By way of example, we obtain

$$\begin{aligned}
\overline{k}_{1,2} &= (K-1)(1-\overline{C}), \\
\overline{k}_{2,3} &= (K-1) - \frac{\overline{C}(1-2q(1-\overline{C})}{1-\overline{C}}\frac{K(K-1)(1-\overline{C})}{(K-1)(1-\overline{C})-1}), \\
&\cdots \quad \cdots \qquad \cdots.
\end{aligned} \tag{19}$$

In general we notice that the following recursion relation is satisfied;

$$A_{i,i+1} = \frac{\overline{C}K(K-1)(n_{i-1}-1) - 2qA_{i-1,i}\overline{k}_{i-1,i}n_{i-1}}{\overline{k}_{i-1,i}(\overline{k}_{i-1,i}-1)n_{i-1}(n_{i-1}-1)}, \tag{20}$$

where

$$A_{i,i+1} \equiv K - 1 - \overline{k}_{i,i+1}. \tag{21}$$

A_i means the number of edges connected between the same generation G_i. We can numerically solve this recursion relation with respect to $\overline{k}_{i-1,i}$.

In order to get numerical value of $\overline{k}_{i-1,i}$, we need to fix three parameters, q, K, and $\overline{C}$. When $\overline{k}_{i-1,i}$ is calculated, N is found from Eq. (9) and (10). If network topology was a tree structure, information would spread over $\sim K^l$ persons at l steps as stated before. In the present case, it is inferred that the clustering coefficient has strong influence on the propagation of information. Then the spread of information would be restricted to rather less persons in networks with large C. We consider how the spread of information is restricted owing to C. We measure it by propagation ratio R that the ratio of the population N really conveyed in the present case to K^l;

$$R = \frac{N}{K^l}. \tag{22}$$

We evaluate R at $l = 6$ with a wide range of the three parameters q, K, and $\overline{C}$ within positive $\overline{k}_{i-1,i}$. They are listed in Table 3.

Table 3. Propagation ratio and C_{min} in parameter space (q, K) at $L = 6$

K/q	q=0	q=0.1	q=0.2	q=0.3	q=0.4	q=0.5
$K = 10^3$	0.996-0.229	0.996-0.032	0.996-0.020	0.996-0.025	0.996-0.028	0.996-0.013
	C =0.29	C =0.33	C =0.38	C =0.42	C =0.46	C =0.51
$K = 10^2$	0.961-0.035	0.961-0.037	0.961-0.025	0.961-0.029	0.961-0.016	0.961-0.018
	C =0.28	C =0.32	C =0.37	C =0.41	C =0.46	C =0.50
$K = 50$	0.876-0.004	0.876-0.041	0.876-0.028	0.876-0.001	0.876-0.001	0.876-0.023
	C =0.27	C =0.31	C =0.36	C =0.4	C =0.46	C =0.49
$K = 28$	0.82-0.018	0.82-0.0015	0.82-0.0054	0.82-0.0102	0.82-0.0147	0.82-0.0175
	C =0.27	C =0.32	C =0.36	C =0.4	C =0.44	C =0.48
$K = 10$	0.664-0.046	0.664-0.041	0.664-0.039	0.664-0.006	0.664-0.006	0.664-0.030
	C =0.21	C =0.25	C =0.29	C =0.33	C =0.39	C =0.42

We find the following facts that by observing in entire parameter region of (q, K)-space;
(1)$0.29 < C < 0.51$,
(2)$N \sim (100 \sim$ a few percent$)$ of $K^{d=6}$,
(3)$\overline{k}_{i,j}^{j} \sim \frac{K}{2}$.

Information can spread over a large portion of K^l, even when C has rather a large value. It is like "small world network" proposed by Watts and Strogatz[2],[3]. From Table 3, we find that a person needs only have about 50 acquaintances in order that information can spread over a few hundred million from only a person even at the worst with the largest clustering coefficient. This condition may be satisfied rather easily in the actual society.

4   Some Empirical Networks

We consider data on Mixi[12] and Bacon Game[17] as suitable networks for our aim.

4.1   Mixi Data

We can interpret Table 4[11] to mean that a person can convey information to how many persons at each generation. First three columns in the Table 4 is based on the data described by Masuda[11]. The last column is average propagation coefficients $\overline{k}_{i-i,i}$ at each generation. Every column in the Table 4 except the first one is shown in Fig. 4 where the length means the distance between a root, that is, a first person and a descendant on the graph. Notice that the length is also the same as generation number i. The distribution of the number of new nodes at G_i is a bell-shaped with the peak at the $length = 5$. Its cumulative distribution has a logistic shape that shows nearly all of the population are included up to the $length = 6$. It drastically rises at the $length = 5 \pm 1$. It is, meanwhile, speculated that some hubs are brought over to the distribution. This suggests that this network is scale free as also suggested by Yuta et al. [12]. The propagation coefficient exhibits a strange behavior at first glance. We speculate that this is an incidental event as will be observed in the next subsection.

Table 4. Data on Mixi by ATR group

generation i	number of new nodes at G_i	number of total nodes	propagation coefficient
0	1	1	28
1	28	29	9.4642857
2	265	294	24.430189
3	6474	6768	13.792864
4	89295	96063	1.9807044
5	176867	272930	0.417692
6	73876	346806	0.167334452
7	12362	359168	0.1173758
8	1451	360619	0.0978635
9	142	360761	0.2112676
10	30	360791	0.133333333
11	4	360795	1.250
12	5	360800	0.4
13	2	360802	
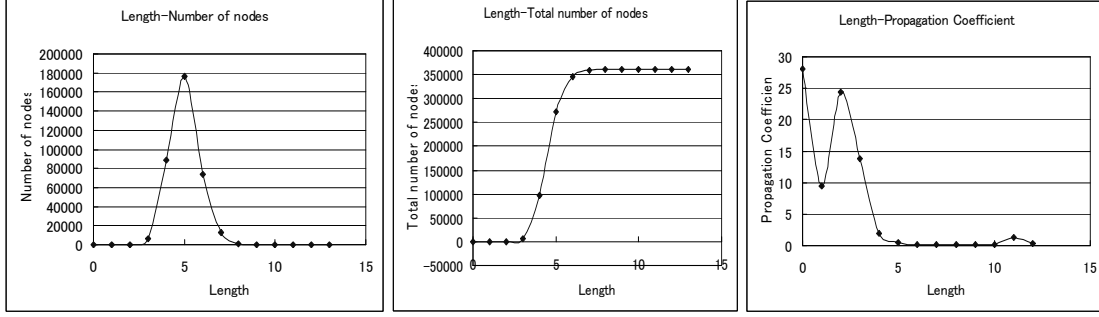
FigD4 Propagation In Mixi. @@

@

Since we can find $\overline{k}_{i-i,i}$ and n_i from Table 4, if K is determined, we can evaluate clustering coefficients by Eq (16) for various values of q. While the average degree is $< K >= 10.4$ in the data for Mixi according to Yuta et al.[12], we assume $< K >= k_{0,1}$ so as to get adequate values for the clustering coefficient. This is the assumption that the number of edges left from a first one person is K. Though needless to say that is not always the case, this assumption can derive meaningful clustering coefficients in the present case. The values of them are given by Table 5. This roughly show the following relation holds;

$$C_i \sim 10^{-(G_i+1)}. \tag{23}$$

When $i > 3$, the structure of the Mixi network is practically a random graph. This also means that C_i is not constant and so the homogeneous hypothesis is rejected in the present case. This is due to almost scale free nature of the Mixi network and finite size effects. Though considerations in the section 2 is not adequate for such cases, they would be adequate to Watts-Strogatz type networks.

Table 5. Clustering Coefficients for Mixi data based on propagation coefficients with $< K >= k_{0,1}$.

|	q=0.5	q=0.8
$G_1 = 1$	none	none
$G_1 = 2$	$9.6 \times 10^{-3\sim4}$	1.4×10^{-3}
$G_1 = 3$	7.6×10^{-5}	8.3×10^{-5}
$G_1 = 4$	7.4×10^{-6}	8.8×10^{-6}
$G_1 = 5$	5.7×10^{-7}	7.5×10^{-7}
$G_1 = 6$	none	1.9×10^{-8}

4.2   Bacon Game

We can draw a good deal of information from the web page[17]. Fig. 5 displays figures similar to Fig.4 in the case of Bacon game. The above figures in Fig.6 are drawn when the starting person is literally Bacon and the below ones are done when the starting person is the one with the longest average path length on the actor network. They are each other quite alike apart from the position of the peak of the distribution (both ends) or maximum gradient (the middle). Key actors would connect with all the actors by short distances so that the distance at the peak is also small. Especially the propagation coefficient marks its peak at the first step for the leading actor Bacon. These observations demonstrate that the homogeneous hypothesis is not valid like the subsection 4.1.

In these experimental networks, we find that C_i does not take a constant value mainly due to finite size effect and almost scale free nature. A simple estimation of C_i occasionally leads to some negative value owing to these. A problem in which C_i becomes negative when we take $K = 10.4$ in the subsection 4.1 crops up. On closer investigation, a finite size effect brings on $\bar{k}_{i-1,i} < 1$ and hubs alone are connected to extraordinarily many persons. These cause negative C_i.

In order to solve the problems, we should consider this subject more minutely beyond the homogeneous hypothesis, including distribution of C_j of the individual, a degree correlation between neighboring persons and so on.

8
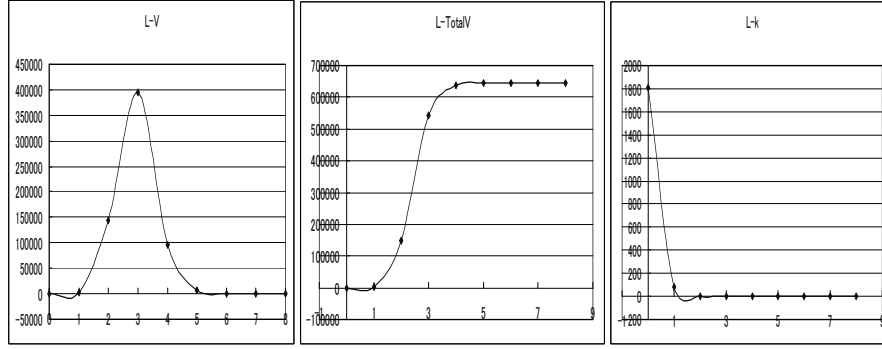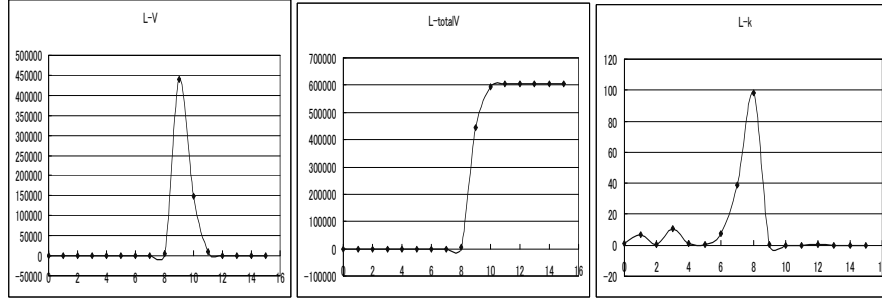
a) Bacon

b) Worst Actor

FigD5 Propagation In Bacon Game. Left: the number of new nodes appeared at a generation G_i. Middle: the number of total nodes appeared up to the generation G_i. Right: The propagation coefficient at the generation G_i. @@

@

5   Summary

In this article we investigate three points about six degrees of separation proposed by Milgram[1]. First one is the analyses of the Watts-Strogatz type small world network, which is based on analytic expressions for the clustering coefficient and the average path length. Second we gave discussions on the propagation coefficient model based on the homogeneous hypothesis for information transmission. Though empirical networks do not necessarily support the hypothesis, we turn out to establish the formalism to calculate propagation coefficient and so on in the schemes where information propagates from generation to generation. Third some experimental networks were investigated where the validity of the homogeneous hypothesis was examined so that some points at issue were made clear. Lastly numerical analyses carried out in these three subjects. Knowledge gained through these investigations is summarized as follows.

1. The effect of the clustering coefficient on diffusion of information over networks of human relations is not so crucial. The effect of clustering coefficient on them only reduces the population who can receive information in a tree graph to a few percentage of it. Though there is a double-digit decrease in the population who can receive information, each person has only to have dozens of contacts for six degrees of separation. Thus "six degrees of separation" is not so amazing phenomenon.

2.The homogeneous hypothesis should be made more accurate. By way of example, considerations for distribution of the clustering coefficient in every node, degree correlation between neighbors as well as degree distribution and so on should be given. They are future issues to be addressed.

3. The finite size effect of networks has to be considered in the analysis of empirical data.

They are summaries and future issues of this article.

References

[1] S. Milgram, "The small world problem", Psychology Today 2, 60-67 (1967)

[2] D. J. Watts@and S. H. Strogatz, "Collective dynamics of 'small-world' networks",@Nature, 393, 440-442(1998)

[3] D. J. Watts, "Six degree– The science of a connected age", W.W. Norton and Company, New York (2003)

[4] R.Albert and A-.L. Barabasi, "Statistical Mechanics of complex networks",Rev. Mod. Phys. 74, 47-97(2002)

[5] A.-L.Barabasi and R.Albert, "Emergence of scaling in random networks", Science, 286, 509-512(1999)

[6] A.-L.Barabasi and R.Albert, "Linked: The New Science of Networks", Perseus Books Group (2002) Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life Plume ; ISBN: 0452284392 ; Reissue (2003/04/29)

[7] P. Erdos and A. Renyi," On random graphs I", Publications Mathematicae Debrecen6, 290-297, 1959

[8] M.E.J. Newman, A.-L.Barabasi and D. J. Watts, "The Structure and Dynamics of Networks", Princeton Univ. Press, 2006 @

[9] S. N. Dorogovtsev, A.V. Goltsev and J.F.F. Mendes, "Pseudo fractal scale-free web", Phys. Rev. E.65, 066122(2002)

[10] S. N. Dorogovtsev and J.F.F. Mendes, "Evolution of Network", Oxford Univ. Press, Oxford(2003)

[11] N. Masuda and N. Konno, "Fukuzatu network towa nanika", in Japanese (Kodanshya 2006)

[12] K. YutaC N.Ono and Y. Fujiwara,"A gap in the community-size distribution of a large-scale social networking site", preprint arXive:physic/0701168.

[13] M.E.J. Newman, "The structure of scientific collaboration networks", Proc. Natl. Acad. Sci. USA 98(2001)404; "A sutudy of scientific coauthorship networks, scientific collaboration networks I, Network consideration and fundamental results", Phys.Rev. E64, (2001)016131; "scientific collaboration networks II,Shortest path, weighted networks, and centrality", Phys.Rev. E64, (2001)016132

[14] A.-L.Barabasi, H. Jeong, Z.Neda, E.Ravasz, A Shubert and T. Vicsek, " Evolution of social network of scientific collaborations", Physica A311, (2002)590

[15] C. Ferrer I and R.V.Sole, "The small world of human language", Proc. R. Soc. London B268(2001)2261

[16] P.A. Fell and A. Wagner, "The small world of metabolism", Nature Biotechnology18(2000)1121

[17] Patrick Reynolds and the CS Web Team," The Oracle of Bacon at Virginia", http://oracleofbacon.org/

[18] A. Barrat and M. Weigt, "On the properties of small-world networks", Eur. Phys. J. B13, (2000)547

[19] M.E.J. Newman, C. Moor and D. J. Watts, "Mean-field solution of small-world network model", Phys. Rev. Lett. 84(14)3201